

## Региональная параметризация глобальной модели растительности SEVER на основе ассимиляции спутниковой карты земного покрова России

С.А. Хвостиков<sup>1</sup>, С.В. Веневский<sup>2</sup>, С.А. Баргалева<sup>1</sup>

<sup>1</sup> *Институт космических исследований РАН, Москва, Россия*

*E-mail: khvostikov@d902.iki.rssi.ru*

<sup>2</sup> *Center for Earth System Science, Tsinghua University, Beijing, People's Republic of China*

В статье описаны метод и результаты параметризации глобальной динамической модели растительности SEVER на территории России с помощью двухэтапной оптимизации ее параметров. Целью параметризации являлось повышение степени сходства между результатами моделирования и спутникового картографирования растительного покрова России. В качестве критерия сходства использовалась оценка, основанная на корреляции пространственного распределения функциональных типов растительности. Для оптимизации и повышения быстродействия имеющейся программной реализации модели SEVER была проведена ее модификация, обеспечивающая, в частности, многопоточные вычисления и асинхронную загрузку данных. Параметризация модели осуществлялась с помощью метода оптимизации EGO (Efficient Global Optimisation), примененного к упрощенной модели, а также квазиньютоновского метода BFGS для исходной модели. В результате оптимизации был получен набор параметров модели, обеспечивающих для территории России более высокий уровень сходства результатов моделирования и спутникового картографирования растительного покрова.

**Ключевые слова:** глобальная модель растительности покровов, параметризация, дистанционное зондирование.

### Введение

Одним из потенциально эффективных инструментов прогноза динамики биосферы в условиях изменений климата являются глобальные динамические модели растительного покрова (DGVM Dynamic Global Vegetation Models), предназначенные для воспроизведения пространственного распределения различных его типов в масштабах планеты (Sitch, 2003; Cox, 2001). Такого рода модели находят все более широкое применение в исследованиях биогеохимических циклов и процессов обмена энергией между наземными экосистемами, атмосферой и другими компонентами геосистемы (Krinner, 2005; Cox, 2000; Cramer, 2001).

Как правило, существующие глобальные модели растительного покрова используют эмпирические оценки значений своих параметров, полученные на основе ограниченных по объему данных локальных экспериментальных исследований, (например, Woodward, 1987), вследствие чего они могут являться неоптимальными.

В последние годы быстрое развитие приобретают методы глобального картографирования растительного покрова на основе данных дистанционного зондирования со спутников (Friedl et al., 2002; Bartholome and Belward, 2005; Defourny et al., 2010). Вместе с тем, несмотря на наличие такого рода глобальных карт, они редко используются для улучшения качества работы моделей растительности (Pfeifer, 2011). В большинстве случаев исследователи ограничиваются сравнением спутниковых карт растительности с результатами моделирования (например, Sitch, 2003), не осуществляя настройку модели по этим данным.

В настоящей работе на основе спутниковой карты растительного покрова России (Барталев и др., 2011) предпринята попытка региональной оптимизации значений параметров глобальной модели SEVER DGVM (Venevsky, 2007), ранее продемонстрировавшей свою принципиальную возможность воспроизведения характеристик растительного покрова на больших территориях.

В силу того что, как правило, глобальные модели описывают процессы динамики растительности с использованием большого количества параметров и уравнений, учитывающих, в том числе, межвидовую конкуренцию, их эмпирическая калибровка путем одновременного варьирования значений нескольких параметров трудно реализуема и может приводить к непредсказуемым результатам. В этой связи, выполненный в настоящей работе подбор наилучших для территории России значений параметров модели, осуществлялся с помощью алгоритмов оптимизации.

### Модель и данные

Глобальная модель растительности SEVER моделирует процессы в десяти функциональных типах растительности, включая хвойные и лиственные леса бореальной, умеренной и тропической климатических зон, тундровую и травянистую растительность на уровне клеток с пространственным разрешением  $0,5^\circ \times 0,5^\circ$  по широте и долготе.

При этом в каждой из клеток независимо моделируется поведение среднего представителя каждого типа растительности, распространяемое на всех его представителей в данной клетке. В начале моделирования SEVER создает популяцию каждого типа растительности во всех клетках и позволяет им развиваться в течение нескольких столетий для достижения равновесного состояния основных характеристик модели. После достижения равновесного состояния осуществляется процесс моделирования динамики растительности в каждой клетке по заданным климатическим данным.

Модель воспроизводит процессы в растительности и окружающей среде, включая фотосинтез, водный обмен, эвапотранспирацию, смерть растений и отмирание отдельных их частей, конкуренцию за свет и некоторые другие. Биомасса растений представлена в модели различными фракциями, такими как листья, корни, физиологически активный слой (заболонь) и внутренняя часть (ядро) ствола для деревьев, а также листья и корни для травянистых типов. Биомасса корней участвует в уравнениях водного баланса и эвапотранспирации, а биомасса листьев необходима для описания процессов фотосинтеза. Кроме того, биомасса различных фракций растений определяет их затраты на дыхание.

Водный баланс в почве моделируется на двух уровнях глубины (0,5 м и 1 м), в каждом из которых доля корней задается для каждого типа растительности. Влажность почвы вычисляется независимо на каждом уровне с учетом осадков, снеготаяния, испарения, поглощения и стока воды. Модель определяет потенциальную эвапотранспирацию, зависящую от интенсивности фотосинтеза, а также температуры и продолжительности светового дня. При этом в случае недостатка влажности почвы и биомассы корней для поддержания эвапотранспирации, растение оказывается в состоянии водного стресса.

Для моделирования процессов фотосинтеза используются данные о входящей радиации, температуре воздуха, продолжительности светового дня, атмосферном давлении, наличии температурного или водного стресса, а по его результатам определяется полная первичная продукция растений. Для получения на ее основе чистой первичной продукции вычитаются затраты на дыхание, зависящие от биомассы различных фракций растений, а также фиксированная ее доля, необходимая для создания новых тканей и размножения.

Чистая первичная продукция используется для увеличения биомассы различных фракций растений, распределяемой между ними с учетом имеющихся для данного типа растительности аллометрических соотношений. Прирост биомассы листвы ведет к увеличению проективного покрытия индивида данного типа и к росту его доли в клетке. Последняя вычисляется как произведение площади проективного покрытия на плотность данного типа, которая возрастает в результате размножения. Если в результате прироста биомассы суммарное проективное покрытие растений превышает размер клетки, то излишки биомассы отмирают с сохранением соотношения прироста биомасс между типами (этот процесс называется конкуренцией за свет).

Часть тканей растения ежегодно отмирает, образуя опад (отмершие части растений на поверхности почвы) или ядро ствола. Также моделируется смертность растительности в результате конкуренции за свет, недостатка чистой продукции, температурного и водного стресса, выход за биоклиматические рамки для данного вида. Умершие деревья добавляются в опад или отмершую биомассу в почве (корни).

Для моделирования используются климатические данные NCEP (<http://www.cpc.ncep.noaa.gov/>) за 1957–2006 гг., интерполированные до пространственного разрешения  $0,5^\circ \times 0,5^\circ$ . Данные о концентрации  $\text{CO}_2$  и почве получены из тех же источников, что и в модели LPJ (Sitch, 2003).

Оценка результатов моделирования выполнялась на основе их сравнения с картой растительного покрова России TerraNorte RLC (Барталев, 2011), полученной на основе спутниковых данных спектрорадиометра MODIS, отражающей современное географическое распределение 23 различных классов земного покрова с пространственным разрешением 230 м.

Сравнение распределения типов растительности по карте TerraNorte RLC и результатов моделирования (*рис. 1а, в*) наглядно демонстрирует различия (в частности, на юго-западе России и на Дальнем Востоке), указывающие, в частности, на потенциальную возможность оптимизации исходных значений параметров модели. Одновременно следует отметить, что отсутствие в текущей версии SEVER возможностей учета антропогенных факторов воздействия на наземные экосистемы (в частности, вырубка лесов и сельскохозяйственная деятельность) также приводит к существенным (особенно в Европейской части России) отличиям результатов моделирования растительного покрова от его современного состояния. Очевидно, что снижение указанных различий, вызванных упомянутым выше ограничением модели, принципиально не может быть обеспечено ее региональной параметризацией.

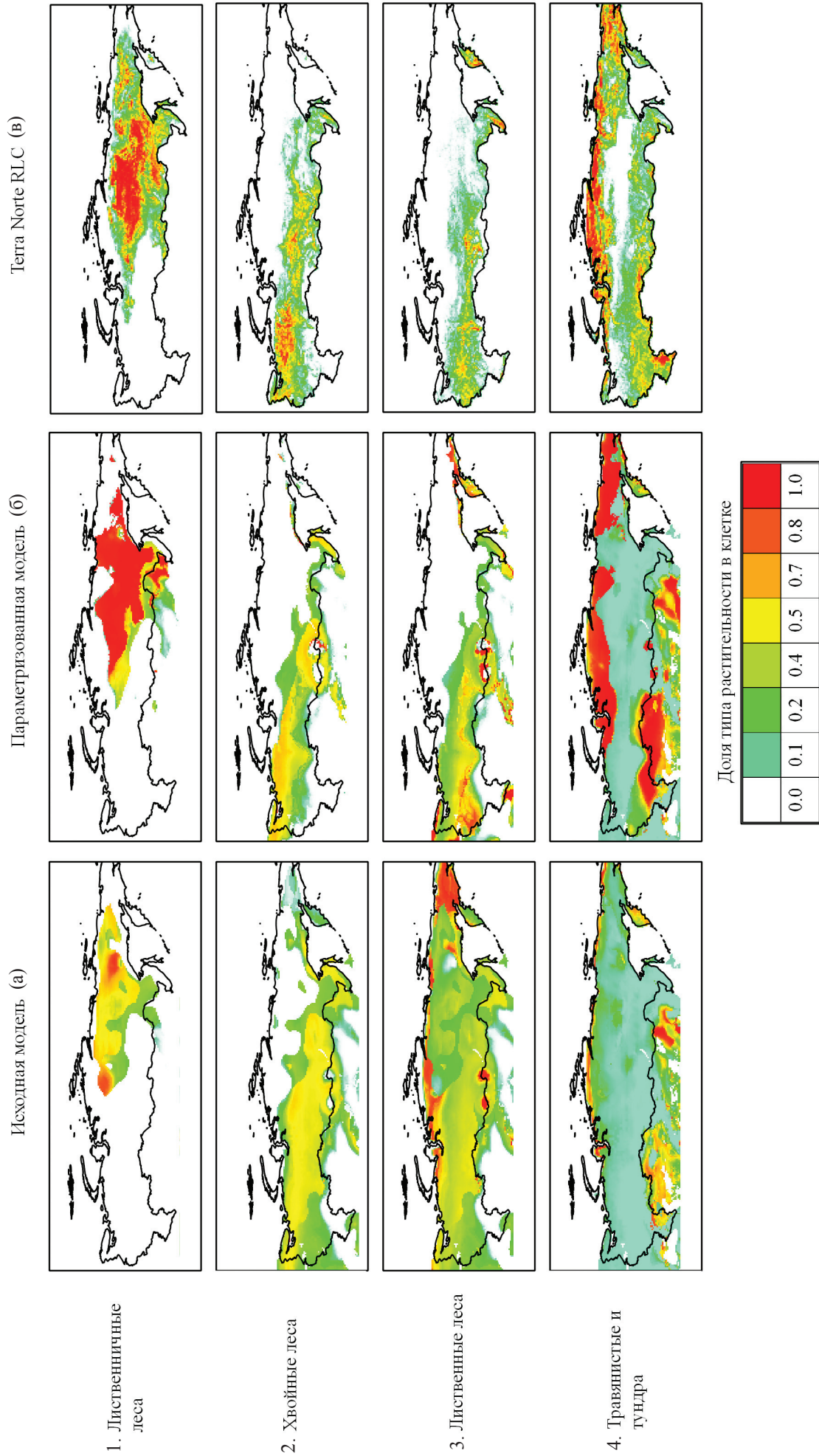


Рис. 1. Сравнение распределения растительности по данным исходной и параметризованной модели и карты Terra Norte RLC

### Критерий качества моделирования

Для оценки точности работы модели осуществлялось сравнение, полученных по результатам моделирования и данным спутниковой карты, значений доли клетки, занятой каждым типом растительности. В качестве критерия качества моделирования использовалась величина, основанная на корреляции между двумя пространственными распределениями значений доли типов растительности, взвешенных, для учета широтных различий размеров клеток, на их площадь, а именно:

$$\bar{Y}_\omega = \frac{\sum_i \omega_i y_i}{\sum_i \omega_i} \quad \text{и} \quad \bar{X}_\omega = \frac{\sum_i \omega_i x_i}{\sum_i \omega_i}$$

$$\text{cov}(X, Y)_\omega = \frac{\sum_i \omega_i (x_i - \bar{X}_\omega)(y_i - \bar{Y}_\omega)}{\sum_i \omega_i}$$

$$\text{corr}(X, Y)_\omega = \frac{\text{cov}(X, Y)_\omega}{\sqrt{\text{cov}(X, X)_\omega \text{cov}(Y, Y)_\omega}},$$

где  $x_i$ ,  $y_i$  – доли типа растительности в  $i$ -ой клетке площадью  $\omega_i$  по карте TerraNorte RLC и по данным модели, соответственно.

Суммирование происходило для клеток с отличными от нуля значениями  $x_i$  или  $y_i$ , а в качестве общей оценки качества моделирования использовалась сумма квадратов корреляций для всех типов растительности.

Как следует из *табл. 1*, оценка результатов моделирования на основе исходных значений параметров демонстрирует для различных типов растительности слабую корреляцию (вплоть до полного ее отсутствия для лиственных лесов), являясь достаточным основанием для проведения исследований по оценке возможностей региональной параметризации модели, как одного из потенциальных путей повышения ее качества.

Таблица 1. Значения коэффициентов корреляции (R) и критерия качества при использовании исходных и оптимизированных параметров модели

<i>Тип растительного покрова</i>	<i>R</i>		
	<i>Исходная модель</i>	<i>Улучшенная модель</i>	<i>Параметризованная модель</i>
Травянистая растительность и тундра	0,33	0,29	0,50
Хвойные вечнозеленые леса	0,46	0,44	0,55
Хвойные листопадные леса	0,36	0,63	0,60
Лиственные леса	-0,04	0,34	0,42
Критерий качества	0,45	0,79	1,10

При визуальном сравнении распределения растительности по результатам моделирования с использованием исходных параметров и данным карты TerraNorte RLC обращают на себя внимание следующие очевидные различия:

- 1) низкий уровень представленности тундры на севере России по данным модели (рис. 1.4а) в сравнении со спутниковой картой (рис. 1.4в);
- 2) недостаток травянистых типов растительности на юго-западе страны по данным модели (рис. 1.4а) в сравнении со спутниковой картой (рис. 1.4в), очевидно связанный с не учитываемой в модели сельскохозяйственной деятельностью;
- 3) низкая представленность лиственных (хвойных листопадных) лесов в Сибири и на Дальнем Востоке по данным модели (рис. 1.1а), в то время как согласно карте TerraNorte RLC (рис. 1.1в) указанный тип растительности преобладает на этой территории;
- 4) невоспроизводимость моделью преобладания хвойных вечнозеленых лесов на северо-западе (рис. 1.2в) и лиственных лесов на юго-западе (рис. 1.3в) страны, в то время как доля указанных типов растительности примерно одинакова на западе России (рис. 1.2а, 1.3а);
- 5) невоспроизводимость моделью особенностей растительного покрова горных регионов Урала, Саян, Среднесибирского плоскогорья (рис. 1а, рис. 1в).

#### *Усовершенствование модели SEVER*

Следует также отметить, что в рамках настоящей работы в модель SEVER были внесены три существенных дополнения, позволивших заметно улучшить качество моделирования.

В частности, в модель была введена дополнительная коррекция температуры воздуха на высоту местности над уровнем моря, основанная на оценке перепада значений высот в клетке и введении отрицательного температурного градиента, равного 0,6°C на 100 м увеличения высоты. Указанная фиксированная величина температурного градиента использовалась в модели в качестве начального приближения данного параметра и в дальнейшем была оптимизирована. Внедрение данной коррекции призвано повысить качество моделирования растительного покрова горных районов.

Кроме того, в модели SEVER было дополнительно учтено влияние высоты растений в их конкуренции за свет. Так, при определении доли растений, умерших в результате конкуренции за свет, ее прирост для каждого типа растительности умножался на коэффициенты вида  $\omega_i = A \frac{(h_i - \bar{h})}{\bar{h}}$ , где А равно 10 (в дальнейшем это значение также было оптимизировано),  $h_i$  – высота среднего индивида для данного типа и  $\bar{h}$  – средняя высота для всех типов растений в клетке. Введение указанных коэффициентов значительно увеличивает долю более высоких растений в клетке и направлено на повышение качества моделирования в части повышения достоверности воспроизведения соотношения хвойных и лиственных лесов.

В качестве еще одного важного усовершенствования в модель SEVER была включена имитация условий произрастания тундровой растительности путем добавления

нечетких ограничений на параметры минимальной температуры, максимальной скорости ветра и максимальной глубины снежного покрова за год, имеющих следующий вид:

$$f(x) = \begin{cases} 0, & x < X_{\min} \\ \frac{(x - X_{\min})}{(X_{\max} - X_{\min})}, & X_{\min} \leq x \leq X_{\max} \\ 1, & x > X_{\max} \end{cases}$$

где  $x$  – параметр, а  $X_{\max}$  и  $X_{\min}$  – указанные в *табл. 2* минимальные и максимальные значения для каждого из параметров.

Таблица 2. Граничные значения параметров окружающей среды для имитации в модели SEVER условий произрастания тундровой растительности

<i>Параметр</i>	<i>Минимальное значение параметра</i>	<i>Максимальное значение параметра</i>
Максимальная глубина снежного покрова, м	0	0,1
Минимальная температура воздуха, °С	-12,2	-35,4
Максимальная скорость ветра, м/с	1,875	5,46

При этом точка относилась к тундре при условии превышения произведения значений  $f(x)$  для всех параметров пороговой величины 0,5. Представленные в *табл. 2* значения параметров были получены путем оптимизации, направленной на минимизацию различий между пространственным распределением тундровой растительности согласно модели SEVER и карте TerraNorte RLC. С учетом приведенного выше усовершенствования, доля клеток с правильным воспроизведением моделью SEVER наличия на территории России тундровой растительности выросла до 60%. Стоит отметить, что верхнее ограничение на глубину снежного покрова оказалось равно 10 см., т.е. условие по глубине снежного покрова свелось к условию наличия снега. Это может быть связано с неточностями в части модели по вычислению глубины снежного покрова.

Совокупность описанных выше усовершенствований модели SEVER позволила повысить качество моделирования (*табл. 1*), в частности, за счет существенного повышения правильности воспроизведения моделью пространственного распределения лиственных и лиственничных лесов. Наблюдаемое при этом снижение корреляции для некоторых типов растительного покрова может объясняться заданием на этом этапе неоптимальных значений параметров модели.

### **Методика параметризации модели SEVER**

Для повышения качества моделирования была проведена региональная оптимизация параметров модели, наиболее существенным образом влияющих на пространственное распределение типов растительности. При выборе подлежащих оптимизации параметров

модели, предпочтение отдавалось отличающимся простотой интерпретации и трудностью непосредственной оценки ввиду нечеткости физического смысла. В модели имеется несколько такого рода параметров, включая, например параметры эмпирических ограничений на температуру воздуха, при которых в растительности могут происходить процессы, необходимые для существования и размножаться данного типа. Параметры уравнений фотосинтеза, эвапотранспирации и водного баланса не включались в процедуру формальной оптимизацию в силу хорошей физической или химической интерпретируемости, возможности непосредственного определения объективными методами, а также ввиду их неочевидного влияния на пространственное распределение типов растительного покрова. Кроме того, при выборе оптимизируемых параметров учитывались результаты предварительно проведенной оценки чувствительности критерия качества моделирования к независимому варьированию их значений.

С учетом приведенных выше соображений для оптимизации было отобрано 12 параметров модели, девять из которых описывают климатические условия для различных типов растительности, при выходе за границы которых они не могут размножаться или погибают. Кроме того, оптимизации подлежали параметры модели, характеризующие площадь листовой поверхности для лиственных лесов, коэффициент в формуле конкуренции за свет и зависимость температуры от высоты над уровнем моря. Два последних параметра были выбраны в связи с тем, что их исходные значения были выбраны произвольно на этапе усовершенствования модели SEVER.

### *Повышение вычислительной производительности модели SEVER*

Необходимость оптимизация параметров модели в широком диапазоне значений потребовала существенного повышения скорости вычислений и, как следствие, модификации программного обеспечения. Дополнительное сокращение вычислительной емкости процедуры оптимизации было достигнуто использованием выборочной версии модели.

В частности, значительное сокращение времени вычислений с использованием реализующей модель SEVER программы была достигнуто за счет включения в нее возможности поддержки многопоточных вычислений. Многопоточность была реализована на основе платформы Net Microsoft (<http://www.microsoft.com/net>) и достигалась распараллеливанием процесса моделирования отдельных клеток на различные потоки в силу полной независимости их обработки. Это, в свою очередь, потребовало перевода исходного кода главной функции программы SEVER с языка программирования FORTRAN на C++. Тестирование многопоточной обработки, выполненное на вычислительной системе с использованием одного, двух и четырех ядерных процессоров, показало практически пропорциональное числу задействованных ядер повышение скорости моделирования. Данные о времени моделирования при разном числе потоков представлены в *табл. 3*.



Таблица 3. Сравнение времени моделирования при различном числе вычислительных потоков с использованием исходной и выборочной модели SEVER

<i>Число вычислительных потоков</i>	<i>Исходная модель SEVER</i>	<i>Выборочная версия модели SEVER</i>
Однопоточные	7 часов	20 минут
Двухпоточные	4 часа	10 минут
Четырехпоточные	2 часа 20 минут	5 минут

Несколько меньшее теоретически возможного (кратное числу ядер системы) повышение скорости моделирования связано затратами времени на загрузку программой данных для каждой клетки. Для дальнейшего увеличения скорости работы в программу была внедрена асинхронная загрузка данных, что позволило загружать данные для последующих клеток во время процесса моделирования. Внедрение асинхронной загрузки данных в многопоточном режиме дало дополнительное сокращение времени моделирования на 5–15%.

#### *Выборочная оценка точности модели SEVER*

Для сокращения времени оптимизации параметров проводилась выборочная оценка точности модели, при которой обрабатывалась небольшая (5%) часть клеток всей территории с оценкой на их основе точности моделирования. Несмотря на то что по результатам нескольких десятков пробных запусков с разными значениями параметров качество модели при полной и выборочной оценке существенно отличалось (примерно на 10%), величина коэффициента детерминации между значениями критерия качества моделирования с использованием двух подходов составила 0,88. Высокий уровень корреляции между значениями критерия качества моделирования при использовании различных подходов может служить обоснованием замены метода оценки точности при исследовании пространства параметров модели. Однако в силу того, что значения критериев для двух подходов различны, оптимальные значения параметров для них могут также существенно различаться, что приводит к необходимости проверки применимости полученных выборочным методом результатов оптимизации к полному набору клеток модели.

#### *Процедура оптимизации параметров модели*

Для оптимизации параметров модели были последовательно использованы два метода. На первом этапе была предпринята попытка поиска глобального оптимума параметров модели с использованием выборочной оценки точности. Затем для модели с полной оценкой точности был осуществлен поиск локального оптимума, с использованием, в качестве исходной точки, найденного на предыдущем шаге глобального оптимума. Для реализации обоих шагов использовался язык R (R Development Core Team, 2010).

На первом шаге оптимизации нахождение глобального оптимума модели выполнялось с использованием метода EGO (Efficient Global Optimisation) (Jones, 1998), реализованного в составе языка R в качестве пакета DiceOptim (Olivier, 2012). Указанный метод сочетает поиск оптимального значения параметров с исследованием всего пространства параметров с использованием метода аппроксимации кригинг (Simpson, 2001), заменяющего аппроксимируемую функцию суммой полиномиальной и случайной компонент. Кригинг в точности интерполирует исходную функцию в точках, по которым построена аппроксимация, и, в силу стохастической природы, может предоставить оценку среднеквадратического отклонения аппроксимации в любой точке пространства параметров.

Сначала алгоритм EGO на основе набора исходных точек строит аппроксимацию, и находит ее оптимум для значения критерия ожидаемого улучшения (Expected Improvement, EI) в виде:

$$EI(x) = (f_{\min} - \hat{y}) \Phi\left(\frac{f_{\min} - \hat{y}}{s}\right) + s \phi\left(\frac{f_{\min} - \hat{y}}{s}\right),$$

где  $f_{\min}$  – минимальное значение функции, найденной на данный момент;  $\hat{y}$  – значение функции в  $x$ ;  $s$  – ошибка (отклонение) аппроксимации в точке  $x$ ;  $\Phi$  и  $\phi$  – нормальная функция распределения и плотность нормального распределения соответственно. Данный критерий увеличивается, если значение функции в точке близко к минимальному, найденному на данный момент, или меньше его (первое слагаемое) и ошибка аппроксимации (отклонение  $s$ ) в новой точке велика (второе слагаемое).

После нахождения точки, дающей максимальное значение критерия ожидаемого улучшения, новая точка проверяется по исходной функции (те есть по выборочной оценке точности модели), и аппроксимирующая функция обновляется. Эти итерации повторяются до тех пор, пока улучшение значения критерия не достигнет предопределенного порогового значения. Такая структура метода делает его быстрым (за счет поиска оптимума по аппроксимации) и эффективным для глобальной оптимизации (из-за специфической формы критерия).

Тем не менее найденный данным методом глобальный оптимум модели по выборочной оценке качества может не являться локальным оптимумом для модели при полной оценке качества, так как различия между ними достаточно велики, как было показано ранее. Поэтому для нахождения локального оптимума был осуществлен второй этап оптимизации, основанный на квазиньютоновском алгоритме BFGS.

Была использована реализация метода BFGS (Broyden, Fletcher, Goldfarb, Shanno) в функции `optim` стандартного пакета `stats` языка R. Метод BFGS – это широко известный квазиньютоновский алгоритм оптимизации, использующий градиент оптимизируемой функции (вычисляемый напрямую) и ее Гессиан (который аппроксимируется по ранее пройденным точкам функции). Этот алгоритм изменяет значения параметров функции в направлении уменьшения градиента, также используя данные о выпуклости (вогнутости) исходной функции (Гессиан). В данном случае информация о градиенте модели

отсутствовала, поэтому была использована конечная аппроксимация градиента. Этот метод оценки градиента для модели требует существенного времени на выполнение, поэтому второй шаг оптимизации включал небольшое число итераций.

## Результаты и заключение

Первый шаг параметризации, выполненный алгоритмом EGO в результате 1253 запусков сокращенной версии модели, подтвердил ранее сделанное предположение о практической невозможности реализации данного этапа на основе полной оценки точности, потребовавшего бы более 2500 часов компьютерного времени. После выполнения данного шага критерий качества моделирования (сумма квадратов пространственных корреляций) увеличился на 38%, а именно с 0,78 до 1,06.

Второй этап параметризации модели, выполненный с использованием полной оценки точности модели методом BFGS, потребовал 161 итерации и позволил достичь локального оптимума с дополнительным повышением значения критерия качества моделирования с 1,06 до 1,10.

Одновременно с повышением значения формального критерия качества моделирования было достигнуто и визуально более близкое распределение воспроизведенных моделью типов растительности к данным спутниковой карты.

Можно отметить, что в результате выполненных усовершенствований модели и ее параметризации достигнуто значительное повышение качества моделирования (*табл. 1*). При визуальном сравнении (*рис. 1б, в*) наглядно видно, что дополненная и параметризованная модель более достоверно воспроизводит географическое распределение тундровой растительности и лиственничных лесов, а также дает более адекватную оценку соотношения между хвойными и лиственными лесами на западе страны. Вместе с тем остаются нерешенными проблемы наличия ошибок моделирования травянистой растительности на юго-западе России и растительного покрова в горных регионах.

Наибольшей (хотя и ожидаемой) проблемой первого этапа оптимизации явилось большое число итераций модели. Кроме того, в качестве проблемы необходимо также отметить поиск точек оптимума критерия ожидаемого улучшения, так как по мере увеличения числа точек, по которым строилась аппроксимация, также растет и время подсчета аппроксимированного значения функции. Таким образом, время поиска оптимума ожидаемого улучшения линейно увеличивалось по ходу оптимизации, и для последних итераций было сравнимо со временем выполнения выборочной оценки точности модели.

Для второго метода основной трудностью являлся подсчет аппроксимации градиента. Использование методов автоматического дифференцирования могло бы значительно увеличить скорость выполнения этого шага, но особенности программы (в частности использование двух языков программирования – C++ и FORTRAN) затруднило использование подобных методов.

Представленная методология может быть применена к оптимизации модели к другим территориям подобного масштаба и также для оптимизации других глобальных

моделей растительности на основе карт, полученных на основе данных дистанционного зондирования.

Созданные к настоящему времени на основе данных дистанционного зондирования карты потенциально могут быть использованы для улучшения качества моделей растительности (Pfeifer, 2011). Вместе с тем возможность применения спутниковых карт для настройки параметров и валидации моделей растительного покрова ограничивается используемой в них концепцией функциональных типов растительности. Функциональные типы растительности объединяют в себе свойства большой группы реальных видов, сгруппированных по климатическим и экологическим признакам с целью уменьшения числа групп растений, обрабатываемых моделями. Выполненные ранее исследования (Poulter, 2011) показали, что использование четырех различных глобальных спутниковых карт земного покрова может приводить к многократным различиям оценки площади отдельных функциональных типов растительности. Тем не менее, пять представленных в России (и других бореальных зонах) типов, показали для разных наборов данных близкие значения площади. В этой связи можно предположить, что выполненная оптимизация позволила получить достаточно устойчивый набор значений параметров, повышающих точность моделирования распределения типов растительности на территории России.

Недостатком данного метода оптимизации является то, что он требует настройки некоторых параметров алгоритма (выбор пороговых значений, частоты перевычисления аппроксимации, вида функции ковариаций для EGO), а также построения и валидации сокращенного варианта модели для глобальной оптимизации. Возможным дальнейшим направлением развития является разработка полностью автоматического метода калибровки моделей, объединяющего методы глобальной оптимизации и ассимиляции данных, как это уже практикуется в гидрологических моделях (Vrugt, 2005).

Созданные в Институте космических исследований Российской академии наук автоматические методы и технологии обработки спутниковых данных (Барталев и Лупян, 2013; Лупян и др., 2011) позволяют строить временные ряды ежегодно обновляемых карт растительного покрова за период с 2000 г. Появление такого рода временных рядов спутниковых карт открывает принципиально новые возможности динамической калибровки и валидации моделей растительного покрова и его динамики в условиях изменений климата, что может рассматриваться в качестве одного из приоритетных направлений дальнейших исследований.

Представленные в настоящей статье результаты получены при финансовой поддержке Microsoft Research.

### Список литературы

1. Барталев С.А. и др. Спутниковое картографирование растительного покрова России по данным спектрорадиометра MODIS // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 4. С. 285–302.

2. *Барталев С.А., Лупян Е.А.* Исследования и разработки ИКИ РАН по развитию методов спутникового мониторинга растительного покрова // Современные проблемы дистанционного зондирования Земли из космоса. 2013. Т. 10. № 1. С.197–214.
3. *Лупян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В., Крашенинникова Ю.С.* Технологии построения информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 1. С. 26–43.
4. *Bartholome E. and Belward A.S.* GLC2000: a new approach to global land cover mapping from Earth observation data // *Int. J. Remote Sens.* 2005. V. 26. P. 1959–1977.
5. *Cox P.M. et al.* Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model // *Nature*, 2000. Vol. 408. No. 6809. P. 184–187.
6. *Cox P.M.* Description of the TRIFFID dynamic global vegetation model // *Hadley Centre Technical Note*. 2001. T. 24. P. 11–16.
7. *Cramer W. et al.* Global response of terrestrial ecosystem structure and function to CO<sub>2</sub> and climate change: results from six dynamic global vegetation models // *Global Change Biology*. 2001. T. 7. No. 4. P. 357–373.
8. *Defourny P. et al.* Accuracy assessment of global land cover maps: lessons learnt from the GlobCover and GlobCorine experiences // *Proceedings of the living planet Symposium*, SP-686. June 2010.
9. *Friedl M.A. et al.* Global land cover mapping from MODIS: Algorithms and early results // *Remote Sensing of Environment*. 2002. Vol. 83. P. 287–302.
10. *Jones D.R., Schonlau M., Welch W.J.* Efficient global optimization of expensive black-box functions // *Journal of Global Optimization*, 1998. Vol. 13. No. 4. P. 455–492.
11. *Krinner G. et al.* A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system // *Global Biogeochemical Cycles*. 2005. Vol. 19. GB1015.
12. *Pfeifer M. et al.* Terrestrial ecosystems from space: a review of earth observation products for macroecology applications // *Global Ecology and Biogeography*. 2011. Vol. 21. Issue 6. P. 603–624.
13. *Poulter B. et al.* Plant functional type mapping for earth system models // *Geoscientific Model Development*. 2011. Vol. 4. No. 4. P. 993–1010.
14. *Rayner P. J et al.* Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS) // *Global Biogeochemical Cycles*. 2005. Vol. 19. GB2026.
15. *Roustant O., Ginsbourger D., Deville Y.* DiceKriging, DiceOptim: two R packages for the analysis of computer experiments by Kriging-based metamodeling and optimization // *Journal of Statistical Software*. 2012. Vol. 51. No. 1. P. 1–55.
16. *Simpson T.W. et al.* Kriging models for global approximation in simulation-based multidisciplinary design optimization. // *AIAA Journal*. 2001. Vol. 39. No. 12. P. 2233–2241.
17. *Sitch S. et al.* Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model // *Global Change Biology*. 2003. Vol. 9. No. 2. P. 161–185.

18. *Venevsky S., Maksyutov S.* SEVER: A modification of the LPJ global dynamic vegetation model for daily time step and parallel computation // *Environmental Modelling & Software*. 2007. Vol. 22. No. 1. P. 104–109.
19. *Vrugt J.A. et al.* Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation // *Water Resources Research*. 2005. Vol. 41. W01017.
20. *Woodward F. I., Williams B. G.* Climate and plant distribution at global and local scales // *Theory and models in vegetation science*. Springer Netherlands, 1987. P. 189–197.
21. *R Development Core Team.* R: A language and environment for statistical computing // R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL <http://www.R-project.org>.

### **Regional parametrisation of Dynamic Global Vegetation Model SEVER based on assimilation of remote sensing data derived land cover map for Russia**

**S.A. Khvostikov<sup>1</sup>, S.V. Venevsky<sup>2</sup>, S.A. Bartalev<sup>1</sup>**

<sup>1</sup> *Space Research Institute of RAS, Moscow, Russia*  
*E-mail: khvostikov@d902.iki.rssi.ru*

<sup>2</sup> *Center for Earth System Science, Tsinghua University, Beijing, People's Republic of China*

This article describes methods and results of parametrisation of Dynamic Global Vegetation Model SEVER on the territory of Russia using two-step optimization procedure. This parametrisation was targeted to increase of similarity between modeling results and remote sensing based land cover map. A similarity criterion was based on spatial correlation for different plant functional types. To perform parametrisation modifications of SEVER program implementation were done, including addition of multithreaded execution and asynchronous data loading. Parametrisation of model was done using optimization method EGO applied to simplified model, and quasi-newton method BFGS applied to full model. As a result of optimization a set of model parameters was obtained, that increases similarity between modeling results and remote sensing land cover map.

**Keywords:** Dynamic Global Vegetation Model, parametrisation, remote sensing.